Molecular characterization of *Glu-B3* locus in wheat cultivars and segregating populations

Araceli Espí, Marta Rodríguez-Quijano, José F. Vázquez, José M. Carrillo, Patricia Giraldo

A B S T R A C T

Bread wheat quality constitutes a key trait for the demands of the baking industry as well as the broad consumer preferences. The role of the low molecular weight glutenin subunits (LMW-GS) with regard to bread quality is so far not well understood owing to their genetic complexity and to the use of different nomenclatures and standards for the LMW-GS assignment by different research groups, which has made difficult the undertaking of association studies between genotypes and bread quality. The development of molecular markers to carry out genetic characterization and allele determination is demanding. Nowadays, the most promising LMW gene marker system is based on PCR and high resolution capillary electrophoresis for the simultaneous analysis of the complete multigene family. The molecular analysis of the bread wheat *Glu-B3* locus in F_2 and $F_{4:6}$ populations expressed the expected one-locus Mendelian segregation pattern, thus validating the suitability of this marker system for the characterization of LMW-GS genes in segregating populations, allowing for the successful undertaking of studies related to bread-making quality. Moreover, the *Glu-B3* allele characterization of standard cultivars with the molecular marker system has revealed its potential as a complementary tool for the allelic determination of this complex multigene family.

1. Introduction

Flour derived from bread wheat (*Triticum aestivum* L., AABBDD, $2n = 6x = 42$) encompasses unique dough viscoelastic properties conferred by prolamins, seed storage proteins, which are further subdivided into gliadins and glutenins according to their solubility in aqueous/alcohol solutions.

Gliadins are monomeric proteins, classified into α, β, γ and ω types based on their mobility in A-PAGE, being mainly related to dough extensibility. ω gliadins are rich in glutamine, proline and phenylalanine forming hydrophobic interactions. α/β and γ gliadins are rich in cysteine and methionine and can thus establish disulfide bonds (Shewry et al., 2003). Gliadin genes (*Gli* loci) are located on the chromosomes of homoeologous groups 1 and 6.

Glutenins are polymeric proteins directly related to the viscoelastic properties of wheat flour. On the basis of their mobility in 1D-SDS-PAGE, they can be classified into two groups: the high molecular weight glutenins (HMW-Gs, covering a molecular range of 70–90 KDa) and the low molecular weight glutenins (LMW-GS, comprising a molecular range of 20–45 KDa). They respectively represent 10% and 40% of the prolamins. Both types of glutenins cross-link to generate the glutenin polymer, so far one of the largest macromolecules described in nature (Wrigley, 1996).

Genes encoding the HMW-Gs are located on the group 1 chromosomes, comprising the *Glu-A1*, *Glu-B1* and *Glu-D1* loci with each locus encoding for two polypeptides, which are designated subunits. The type and the number of subunits, between 3 and 5, typify the allelic variation of the HMW-GS subunits and are the most important determinants with regard to the rheological quality of bread wheat dough (Payne et al., 1987), therefore having been deeply studied at a protein level. The allelic characterization is easy to undertake by protein electrophoresis, yet in some cases it might be necessary to resort to other techniques comprising a higher resolution such as RP-HPLC, CE or MALDI-TOF-MS (Gao et al., 2010). In addition, molecular marker use is becoming progressively more frequent. PCR based specific markers for loci *Glu-A1*, *Glu-B1* and *Glu-D1* exist (i.e: Ahmad, 2000; Ma et al., 2003). The extended

knowledge base with regard to these proteins has substantially contributed to accelerate the genetic improvement of the bread quality trait in breeding programs.

Compared to the HMW-GS, the LMW-GS are a much more complex family, substantially more LMW-GSs are expressed in the kernels of bread wheat and in consequence their contribution to wheat bread quality characteristics is still ill defined.

Based on the first amino-acid of the N-terminal protein, LMW-GS have been classified as LMW-m (methionine), LMW-s (serine) and LMW-i (isoleucine) (D'Ovidio and Masci, 2004). The proposed primary structure of a typical LMW-GS is: a signal peptide of 20 amino acids followed by a short N-terminal domain (13 amino acids), next to a central repetitive domain and concluding with a C-terminal domain, which is further subdivided into three distinctive regions namely, C-terminal I, II and III (D'Ovidio and Masci, 2004). Whereas the signal peptides and the C-terminal I and III regions are extremely well conserved, the repetitive domains are highly polymorphic in length (D'Ovidio and Masci, 2004).

Most of the LMW-GS genes are encoded by the *Glu-A3, Glu-B3* and *Glu-D3* loci, located on the group 1 chromosomes and closely linked to the γ and ω gliadin loci (Gupta and Shepherd, 1990). Part of the bread quality influence attributed to gliadins can be in fact due to the tightly linked LMW-GS loci. Together these genes comprise a multi-gene family whose number has been estimated to vary from 10 to 40 in hexaploid wheat varieties (Huang and Cloutier, 2008). However, the exact copy number of LMW-GS genes is still unknown, mostly due to a lack of efficient methods to distinguish between members of this multigene family. Studies using cDNA and screening of bacterial artificial chromosome (BAC) libraries (Ikeda et al., 2006; Huang and Cloutier, 2008; Dong et al., 2010) have provided a general elucidation of the complexity regarding members of the LMW-GS gene family in wheat varieties. At least four genes coding mostly types LMW-GS-m and i have been identified for the *Glu-A3* locus (from *A3-1* to *A3-4*); three coding mostly type LMW-GS-s for the *Glu-B3* locus (from *B3-1* to *B3-3*) and seven coding only type LMW-GS-m for the *Glu-D3* locus (from *D3-1* to *D3-7*), nonetheless, the number of genes can vary among cultivars and presumably could be higher in view of the fact that a number of genes of this family are still unknown (Dong et al., 2010). In addition, the presence of pseudogenes has also been described for all loci yet their type and number also depend on the cultivar analyzed.

Usage of molecular markers to discriminate different LMW-GS genes has also been investigated by several groups (Dong et al., 2010; Huang and Cloutier, 2008; Ikeda et al., 2006; Lan et al., 2013; Liu et al., 2010; Long et al., 2005; Sharma et al., 2013; Wang et al., 2009,2010; Zhao et al., 2007a,b). Although some of these markers might be useful in breeding programs, the high LMW-GS gene homology and their high content of repetitive domains have made the molecular approach arduous and most molecular markers proposed so far are not very robust, reducing their wide application.

Nonetheless, recently a new molecular marker method was proposed by Zhang et al. (2011a). This method is not locus or allele specific, the ultimate goal is the complete amplification by PCR of all of the LMW-GS coding sequences of a given cultivar followed by their analysis by Capillary Electrophoresis (CE). This is achieved by using a pool of primers designed according to the conserved regions of the LMW gene sequences. So far, this has been revealed as a very powerful marker system to elucidate the complex members of the LMW gene family in different cultivars and isogenic lines (Zhang et al., 2012, 2013).

The detailed characterization of LMW-GS members is important for accurate bread quality studies since the allelic HMW-GS variation is not sufficient to explain all of the variation observed with regard to the bread quality parameter values. The function of the LMW-GS proteins respective to the control of the end use qualities of wheat kernels has been studied in both tetraploid and hexaploid cultivars; the genetic studies having generally found positive contributions of the *Glu-3* loci in connection with parameters related to dough strength, extensibility and bread-making quality (Békés et al., 2006; Gupta and McRitchie, 1994; Ma et al., 2005; Oury et al., 2010; Payne et al., 1987; Zhang et al., 2012). Nowadays, the molecular mechanisms underlying the functional genetic differences of either the orthologous or the allelic *Glu-3* loci have still not been well investigated. The further development and use of molecular markers which support an analysis of the LMW-GS role concerning bread quality is essential while resorting to populations of advanced bread wheat lines. In particular, the detailed characterization of the allelic diversity present at the *Glu-B3* locus exhibiting the highest complexity and variation is important, as it has been established that the most important contribution to bread wheat quality derives from the LMW-GS loci (Zhang et al., 2012).

The objective of this study has been to gain insight into the molecular characterization of the bread wheat *Glu-B3* locus using a molecular marker approach, concomitantly validating the molecular markers in segregating populations and evaluating their possible role in reference to supporting *Glu-B3* specific allele determination in standard cultivars.

2. Material and methods

2.1. Plant material

Two cultivars were selected for having the same composition at the *Glu-A3* and *Glu-D3* loci, but being different for the *Glu-B3* locus: 'Gazul' a hard-grained spring wheat with a higher gluten strength and 'Tigre' a hard-grained winter wheat with a high dough extensibility. Two populations derived from the 'Tigre' × 'Gazul' cross were used for this work, an F_2 population ($N = 95$) and an advanced $F_{4:6}$ population ($N = 65$). Lines were raised at the experimental field of the School of Agricultural Engineering, Technical University of Madrid, Spain (40°26′47.36″ N, 3°44′21.00″ W).

LMW-GS standard cultivars 'Capelle-Desprez', 'Chinese Spring', 'Fengmai 27', 'Gabo', 'Halberd', 'Heilo', 'Nanbu-komugi', 'Norin 61', 'Pepital' and 'Thesee' were selected according to Liu et al. (2010).

2.2. Protein electrophoresis and nomenclature

Prolamin extraction was performed according to Singh et al. (1991). Allelic variation of LMW-GS was determined using Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (1D-SDS-PAGE) as previously described (Payne et al., 1987). LMW-GS were named following Gupta and Shepherd (1990), Jackson et al. (1996), Igrejas et al. (1999) and Liu et al. (2010).

2.3. DNA and RNA isolation and PCR amplification

Genomic DNA was extracted from seeds using the cetyl trimethyl ammonium bromide (CTAB) procedure (Saghai-Maroof et al., 1984). Total RNA samples were prepared from developing seeds (10 days post-anthesis) using the TRIzol reagent according to the manufacturer's protocol (Invitrogen, Carlsbad, CA) and were converted into cDNA using the PrimeScript™ 1st strand cDNA Synthesis Kit (Takara).

The primer sets used for the LMW-GS amplification were modified from Zhang et al. (2011a). To optimize the different amplification, sets of degenerated primers were designed (Table 1). The forward primer of each pair was labeled with

Table 1
Degenerated oligonucleotides optimized for the amplification of LMW-GS genes. **Y** = C/T, ***R*** = A/G. Primers labeled with 6-FAM are marked with*. Primer Low1-2F* comprised the same forward primer for the Low1 and Low2 primer combinations.

Primer	Sequence (5'-3')
Low1-2F*	ATG AAG ACC TTC CTC **R**TC TTT GC
Low1R	CA**Y** GGG TTT AG**Y** TGC TGC A
Low2R	CAA CAT TGT **Y**G**Y** **Y**GC ATC ACA T
Low3F*	CCT GGT TTG GAG A**R**A CCA TG
Low3R	CAT **Y**TG **Y**GA CCT AGC AAG A**Y**G
Low4F*	CCT **R**G**Y** TTG GAG A**R**A CCA T**Y**G C
Low4R	CAT TTG **Y**GA CCT AGC AAG ACT
Low5F*	GCC GTT GCG CAA ATT TCA C
Low5R	CAT TTG TGA CCT AGC AAG ACA TC

6-carboxyfluorescein (6-FAM) at the 5' terminal end. PCR reactions were performed in 25 µl reaction volumes containing 100 ng of genomic DNA or cDNA, 1x Certamp Complex Buffer (Biotools), 2 µM $MgCl_2$, 200 µM dNTPs, 0.5 µM primers and 1 µl of Certamp Complex Enzyme Mix (Biotools). Amplifications were performed in a MyCycler Thermocycler (BioRad) with the following conditions programmed: 2min at 94 °C, 35 cycles of 30 s at 94 °C, 1 min at 58 °C, and 1 min at 72 °C, and a final extension step of 10min at 72 °C. Results were visualized in 2% agarose gels stained with GelRed (Biotium).

2.4. Capillary electrophoresis analysis (CE)

PCR products with the fluorescent label were diluted 1:10 in water, mixed with Hi Di formamide and the 1200-Liz™ internal size standard, denatured at 95 °C during 5 min and genotyped using an ABI PRISM® 3100 DNA analyzer (Applied Biosystems, USA). Allele size determination of the different DNA fragments was carried out by means of the GeneMarker software v3.7 (SoftGenetics, LLC, USA).

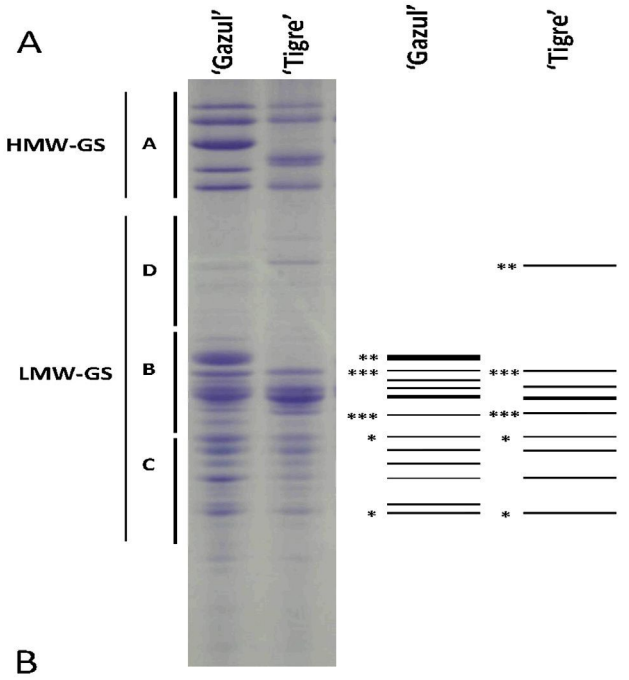
3. Results

3.1. LMW-GS characterization of parental lines

The LMW-GS subunits of the parental lines detected by 1D-SDS-PAGE are shown in Fig. 1. Parental lines presented the same allele for the *Glu-A3* and *Glu-D3* loci, in either case the *b* allele. Protein subunits corresponding to the genetic locus *Glu-B3* were assigned the *d* allele in the case of 'Gazul', compared to the *f* or *g* alleles in the instance of 'Tigre' following the nomenclature of Gupta and Shepherd (1990), Jackson et al. (1996) and Igrejas et al. (1999). Using the standards for the *Glu-B3* alleles derived from Liu et al. (2010), 'Tigre' was allocated the *g* allele while 'Gazul' the *i* allele. Overall, the *Glu-B3* alleles were difficult to uncover by protein electrophoresis, consequently the molecular marker method was applied.

3.2. Molecular marker analysis of segregating populations

The LMW-GS genes of 'Gazul' and 'Tigre' were amplified with the primer pairs Low1, Low2, Low3, Low4 and Low5. These primers anneal to the conserved regions of the LMW-GS and amplify the repetitive domain which holds most of the observed variability (Fig. 2A). The different primer sets are supposed to amplify all of the members of the LMW-GS family including the m-, s- and i-type members. The exact annealing positions of primers to the different LMW-GS sequences available in the GenBank database are shown in Fig. S1.



	Glu-A3	*Glu-B3* (1)	*Glu-B3* (2)	*Glu-D3*
'Gazul'	*b*	*d*	*i*	*b*
'Tigre'	*b*	*f/g*	*g*	*b*

Fig. 1. LMW-GS characterization of parental lines 'Gazul' and 'Tigre'. (A) LMW-GS fractionation by SDS-PAGE electrophoresis. To the right, a schematic pattern of the protein subunits used for the allele determinations is shown. Bands marked with * belong to the *Glu-A3* locus, Bands marked with ** belong to the *Glu-B3* locus, Bands marked with *** belong to the *Glu-D3 locus*. Unmarked bands correspond to protein subunits whose locus assignment has not been determined. (B) Allele assignments of the *Glu-A3*, *Glu-B3* and *Glu-D3* loci. For the *Glu-B3* locus, two different classifications are exemplified: (1) stands for the nomenclature used by Gupta and Sheperd (1990), Jackson et al. (1996) and Igrejas et al. (1999) whereas (2) stands for the nomenclature used by Liu et al. (2010).

As expected, the amplification products as visualized in agarose gels were made up of complex fragment mixtures allowing for the detection of differences between cultivars (Fig. 2B). The CE analysis yielded a lot of fragments between 200 and 700bp appearing in each amplification reaction. Concerning the Low3 and Low5 amplifications both parental cultivars presented the same peak profiles, nonetheless with reference to the Low1, Low2 and Low4 amplifications, a total of twenty-four polymorphic fragments were observed, respectively six, ten and eight (Table 2). The detected polymorphisms reflect differences with regard to the LMW-GS multigene constitution among cultivars, nevertheless, in addition to active genes the fact could also correspond to the presence of pseudogenes. To verify this possibility, amplifications were also performed using cDNA as a template. Results revealed a substantial reduction in the number of peaks (twenty fragments disappeared in the instance of 'Gazul' while sixteen in the case of 'Tigre'), remaining sixteen polymorphic peaks indicative of differences which belonged to active genes (Table 2). A comparison of the size of the DNA fragments obtained in this study with that of previous works (Zhang et al., 2013) made putative loci assignments possible, especially in the case of the Low1 and Low 2 amplifications (Table 2).

The LMW-GS analysis by 1D-SDS-PAGE had previously shown that cultivars 'Gazul' and 'Tigre' enclosed the same composition for

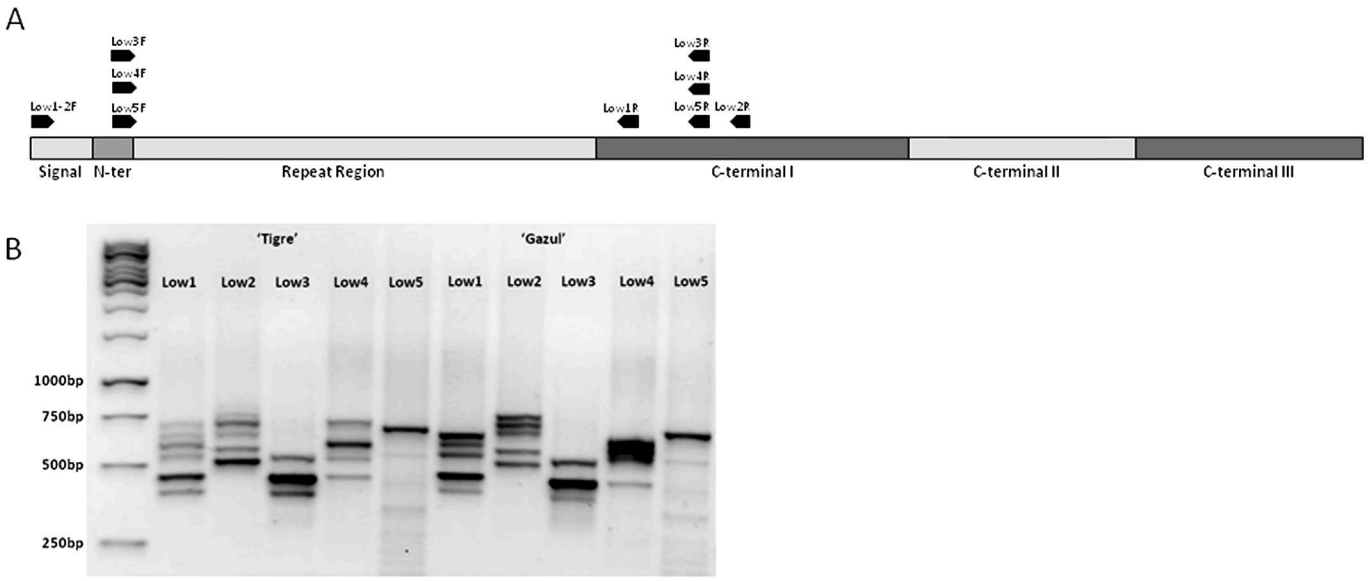


Fig. 2. LMW-GS gene amplifications. (A) Standard LMW-GS structure with indication of the primer locations. LMW-GS is composed of a signal peptide followed by a short N-terminal domain, a central repetitive domain and a C-terminal domain, which is further subdivided into three distinct regions, namely C-terminal I, II and III. Primers were designed flanking the repeated region. (B) PCR amplification with oligonucleotide pairs Low1, Low2, Low3, Low4 and Low5 of cultivars 'Gazul' and 'Tigre'. The first lane displays the 1 Kb DNA marker ladder. Polymorphic differences can perfectly be discerned at this level between cultivars and among primer sets.

the *Glu-A3* and *Glu-D3* loci, nevertheless, displayed differences regarding the *Glu-B3* locus, in this way, the molecular polymorphic fragments presumably could correspond to the *Glu-B3* locus. As expected, no polymorphic fragments were detected for the Low3 and Low5 amplifications considering that these primer pairs should not amplify the *Glu-B3* locus as predicted by the sequence analysis (see Fig. S1).

The segregation of molecular markers was studied in the F_2 population. The DNA derived from 95 individuals was amplified with the Low1, Low2 and Low4 primer sets and subsequently analyzed by CE. For each F_2 individual, the sixteen polymorphic fragments were either scored as present (1) or absent (0) and named according to the primer set number followed by the length of the fragment in base pairs (i.e. 2-616.3 stands for a fragment of 616.3 bp amplified by the Low2 primers). Results are displayed in Table 3. The polymorphic fragments were classified into two groups as presented in Table 3 for the parental cultivars 'Tigre' or 'Gazul'. In a one locus F_2 segregation, a total of 3 different genotypes in a 1:2:1 distribution are expected. Accordingly, in this case three major genotypes were found, two of them corresponding to the homozygous parental genotypes and representing around 25% of the samples and a third class counting up to 50% of the lines, as expected from a monogenic heterozygous frequency and containing all of the polymorphic fragments. The observed F_2 segregation is in agreement with the segregation pattern of one locus with two codominant alleles, supporting the presence of only one segregating locus in the parental lines. Only one individual showed an unexpected fourth genotype, which could reflect an intra-locus recombination event, since all fragments are supposed to belong to the *Glu-B3* locus.

Analysis by means of the molecular approach was also carried out on a population of advanced $F_{4:6}$ lines. DNA obtained from sixty-five lines was amplified with the Low1, Low2 and Low4 primer sets. The polymorphic fragments segregated in the expected 1:1 fashion; half of the lines presented the genotype of 'Gazul' while the other half displayed the genotype of 'Tigre' (Table 3). No heterozygotes were detected; yet more unexpected genotypes than in F_2 were identified (six out of 65 individuals with three unexpected genotypes and presumably corresponding to intra-locus recombination events).

3.3. *Glu-B3 allele determination with the molecular marker method*

Twenty-eight cultivars whose allele composition had previously been determined were selected from the study of Liu et al. (2010), the set included ten cultivars established as standards for the *Glu-B3* alleles. We were not able to obtain the recommended standard cultivars for *the Glu-B3 h* and *Glu-B3 j* alleles; nonetheless, for this purpose, the 'Neepawa' and 'Apollo' varieties were used. DNA from all samples was amplified with the Low2 primer set and further analyzed by CE. Amplified fragments of each cultivar were scored as either present (1) or absent (0) (Table S1). The number of amplified fragments recovered from each cultivar varied between 9 for 'Katepwa' to 24 for 'Chinese Spring'. Some fragments amplified frequently, being present in all (491.7, 654.3, 680.8) or almost all (636.5, 499.0) cultivars. In contrast, others were rarely observed as for instance 698.6 and 729.0 only present in the case of 'Chinese Spring' or 706.8 only appearing for cultivar 'Nambu-komugi'. Noteworthy remains the fact that the universal fragments 491.7, 654.3, and 680.8 presumably consist of pseudogenes, according to this study and also to previous research (Zhang et al., 2013). Nonetheless, PCR amplification is not locus specific, therefore fragments could have corresponded to more or less any of the *Glu-A3*, *Glu-B3* or *Glu-D3* loci. Putative assignments were made, when possible, based on previous research (Zhang et al., 2013) (Table S1).

Standard cultivars were analyzed looking for genotype differences. We focused on a series of polymorphic fragments (616.3, 636.5, 649.8, 675.5, 696.7, 698.6, 706.8, 712.3, 729.0, 791.9 and 794.8) previously assigned to the *Glu-B3* locus by Zhang et al. (2013). Genotype comparisons confirmed that the different standard varieties evidenced different genotypes which could reflect their different allele compositions (Table 4).

The genotypes of the cultivars described with the same *Glu-B3* allele as the standards were then compared, looking for similarities. In general, the cultivars sharing the same allele classification also displayed the same genotype (Table 4). Although the standard

Table 2
Amplified DNA fragment sizes (in bp) of the parental varieties using primers Low1, Low2, Low3, Low4 and Low5. Putative locus assignments according to Zhang et al. (2013) are displayed in the second column.

	Locus	Genomic DNA		cDNA	
		'Tigre'	'Gazul'	'Tigre'	'Gazul'
Low1	*Glu-D3*	385.1	385.1	385.1	385.1
	Glu-A3	391.5[b,c]	391.5[b,c]	–	–
	Glu-D3	392.8[b,c]	392.8[b,c]	–	–
	Glu-D3	394.2[b]	394.2[b]	–	–
	Glu-A3	400.1[b,c]	400.1[b,c]	–	–
	Glu-D3	441.3	441.3	441.3	441.3
	Glu-B3	–	510.2[a]	–	510[a]
	Glu-D3	525.2	525.2	525.2	525.2
	Glu-B3	530.5[a]	–	530.5[a]	–
	Glu-B3	544.0[a]	–	544.0[a]	–
	Glu-B3	548.6[b,c]	548.6[b,c]	–	–
	Glu-B3	–	570.1[a]	–	570.1[a]
	Glu-D3	574.9	574.9	574.9	574.9
	Glu-D3	577.6	577.6	577.6	577.6
	Glu-D3	591.0	591.0[b]	591.0	–
	Glu-A3	620.2	620.2	620.2	620.2
	Glu-B3	–	688.0[a]	–	688.0[a]
	Glu-B3	–	690.5[a]	–	690.5[a]
Low2	–	–	472.8[b]	–	–
	Glu-A3	484.5	484.5	484.5	484.5
	Glu-D3	491.7	491.7	491.7	491.7
	–	–	496.2[b]	–	–
	Glu-D3	499.0[b]	499.0[b]	–	–
	Glu-D3	500.6[b]	500.6[b]	–	–
	Glu-D3	–	503.0[b]	–	–
	Glu-A3	505.9	505.9[b]	505.9	–
	Glu-D3	546.6	546.6	546.6	546.6
	Glu-B3	–	616.3[a]	–	616.3[a]
	Glu-D3	631.6[b]	631.6[b]	–	–
	Glu-B3	636.5[a]	–	636.5[a]	–
	Glu-B3	649.8[b]	–	–	–
	Glu-B3	654.3[b]	654.3[b]	–	–
	Glu-B3	–	675.5[a]	–	676[a]
	Glu-D3	680.8	680.8	680.8	680.8
	Glu-D3	683.1	683.1	683.1	683.1
	Glu-D3	692.0[b]	–	–	–
	Glu-D3	–	696.7[b]	–	–
	Glu-A3	725.2	725.2	725.2	725.2
	Glu-B3	–	791.9[a]	–	791.9[a]
Low3	–	366.6[b]	366.6[b]	–	–
	–	373.7	373.7	373.7	373.7
	–	381.4[b]	381.4[b]	–	–
	Glu-D3	383.2[b]	383.2[b]	–	–
	Glu-A3	388.7	388.7	388.7	388.7
	–	429.9	429.9	429.9	429.9
	Glu-A3	514.6	514.6	514.6	514.6
Low4	–	–	499.6[a]	–	499.6
	–	519.9[a]	–	519.9	–
	–	533.6[a]	–	533.6	–
	Glu-A3	537.7[b]	537.7[b]	–	–
	–	563.9	563.9	563.9	563.9
	–	566.8	566.8	566.8	566.8
	–	575.8[a,b]	–	–	–
	–	–	581.6[a,b]	–	–
	-	610.2[a]	–	610.2	–
	–	–	677.8[a]	–	677.8
	–	–	680.6[a]	–	680.6
Low5	–	523.8	523.8	523.8	523.8
	–	625.0	625.0	625.0	625.0

[a] Polymorphic fragments corresponding to active genes are marked with a.
[b] Pseudogenes identified in this analysis are marked with b.
[c] Pseudogenes identified in a previous study (Zhang et al., 2013) are marked with c.

cultivars for *Glu-B3a*, *Glu-B3f* and *Glu-B3ab* presented differential genotypes, they could not be compared with more cultivars. Cultivars 'ACA601' and 'Marquis' categorized with the *b* allele shared with the standard cultivar 'Gabo' a genotype characterized by the presence of fragments comprising lengths of 636.5 and 712.3. 'Insignia', classified with the *c* allele, manifested the same genotype as 'Halberd'. 'Glenlea' and 'Tigre' presented the same genotype as 'Capelle-Desprez' implying the presence of a *g* allele in 'Tigre'. 'Buck Brasil', 'Klein Proteo' and 'ACA801' previously had been classified as comprising the *g* or *ac* alleles. Nevertheless, in this comparative analysis, cultivar 'ACA801' was classified as *ac*, however 'Buck Brasil' and 'Klein Proteo' exhibited a genotype different from both alleles yet more similar to the *g* allele.

Only the *Glu-B3h* and *Glu-B3d* alleles were indistinguishable. On the one hand, 'Orca' classified as expressing the *d* or *i* alleles comprised a genotype identical to the *h/d* genotype, accordingly could presumably be classified as *d*. On the other hand, 'Chopin' classified as *h* provided a different genotype, with the presence of a 696.7 fragment which corresponds to the *c* allele. The 'Klein Martillo' cultivar, classified on the basis of the *Glu-B3j* allele, did not reveal any fragment in common with the 'Apollo' cultivar which was taken as the *Glu-B3j* standard. Cultivars 'Opata', 'Etoile de Choisy' and 'Gazul' presented genotypes very similar to the standard cultivars for the *ad* allele.

4. Discussion

Usage of different nomenclatures and standards with regard to the assignment of LMW-GS genotypes by different groups has not only made the classification difficult, in addition, has also hampered to undertake association studies between genotypes and bread quality. It is therefore necessary to establish a unique nomenclature and use specific standards for each allele as has already been proposed by Liu et al. (2010). The before mentioned authors resorted to different techniques to determine different alleles, namely, 1D-SDS-PAGE, 2DE, MALDI-TOF and PCR, nonetheless, no single technique was capable of discriminating between all allelic variants. The molecular marker method developed by Zhang et al. (2011a) has been proposed as a powerful method for LMW-GS gene characterization. It contemplates the advantage that with few PCR amplifications, all of the genes are simultaneously amplified, not corresponding to an allele specific amplification so that the detection of new alleles is possible. This is especially important as for instance in the case of the *Glu-B3* locus which comprises a lot of presumably uncharacterized variability.

In comparison to the 2DE analysis, a PCR based marker approach is much easier to execute for any laboratory and does not need any special equipment or expertise as in the case of 2DE; in addition, fragment analysis is normally externalized and constantly becoming cheaper. With regard to the analysis of segregating populations the molecular marker method is much more efficient since a high number of samples can be simultaneously analyzed.

The oligonucleotides employed in this study have been based on those described by Zhang et al. (2011a) albeit simplified in number by using degenerate bases. The five primer pairs Low1, Low2, Low3, Low4 and Low5 performed as expected and the amplification patterns observed have been comparable to those previously described. For amplifications carried out with Low1, the genotype retrieved for 'Gazul' (1-510.2, 1-548.6, 1-570.1, 1-688.0, 1-690.5) was identical to the second most frequent genotype described by Zhang et al. (2013) who observed 38 out of 237 cultivars (including Chinese wheat landraces, modern cultivars and foreign cultivars) with the same genotype. The genotype found for 'Tigre' (1-530.5, 1-544.0, 1-548.6) could have corresponded to the third most frequent genotype described by Zhang et al. (2013) who determined that 24 out of 237 cultivars (including Chinese wheat landraces, modern cultivars and foreign cultivars) displayed a very similar genotype, encompassing fragment 1-544.0 (exclusively) in addition also to a 1-621.1 bp fragment not present in 'Tigre'.

The fact that no polymorphic fragments were observed with primers Low3 and Low5 supports the results of the 1D-SDS-PAGE

Table 3
Presence/absence matrix of polymorphic fragments in the F_2 and $F_{4:6}$ populations. Parental cultivars and the number and percentage of individuals of each population representing genotypes are indicated in the first column.

	1-510.2	1-570.1	1-688.0	1-690.5	2-616.3	2-675.5	2-791.9	4-499.6	4-677.8	4-680.6	1-530.5	1-544.0	2-636.5	4-519.9	4-533.6	4-610.2
'Tigre'	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1
'Gazul'	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0
F_2																
47 (49.47%)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
26 (27.37%)	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0
21 (22.11%)	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1
1 (1.05%)	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1
$F_{4:6}$																
30 (46.15%)	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1
29 (44.62%)	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0
4 (6.15%)	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	1
1 (1.54%)	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	1
1 (1.54%)	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1

analysis which revealed only polymorphism between the parental lines for the *Glu-B3* locus, considering that these primers should not amplify this locus (see Fig. S1).

Pseudogenes of the *Glu-3* loci have been described as very frequent (Dong et al., 2010). One disadvantage of the molecular marker approach consists of the incapability to distinguish between active and inactive genes. In order to simplify the analysis of segregating populations, pseudogenes must be identified in any study as they do not contribute to the final composition of proteins. Pseudogene discrimination can only be achieved by amplifying cDNA. In this study, a DNA vs. cDNA amplification pattern comparison revealed the presence of several pseudogenes belonging to the three *Glu-3* loci. Noteworthy was that one fragment (2-505.9) corresponded to an active gene in 'Tigre' and to an inactive gene in 'Gazul' (Table 2). Similar situations had previously already been depicted by Dong et al. (2010) and Zhang et al. (2011a). Although working with RNA may sometimes be difficult, this type of analysis has only to be addressed using parental lines; once the fragments corresponding to the active genes have been identified, the analysis in segregating populations can be based only on DNA.

To the best of our knowledge, this is the first time that this molecular system has been analyzed in segregating populations, since previous studies have used cultivars or NILs (Zhang et al. (2011a), 2011b, 2012, 2013). The segregation of polymorphic fragments in the F_2 and $F_{4:6}$ generations has confirmed the existence of only one segregating locus.

Cultivar 'Gazul' presented the haplotypes described by Zhang et al. (2013) of *B3-510/570* and *B3-688/691* (referring to amplification with Low1) having conserved these haplotypes from the F_2 to the $F_{4:6}$ generations. The combination of both haplotypes was

Table 4
Genotypes identified by means of Low2 amplifications using cultivars with different *Glu-B3* compositions. Only polymorphic fragments related to the *Glu-B3* locus are shown. Cultivars recommended as standards for the determination of LMW-GS alleles are underlined. Allele assignments by a previous study (Liu et al., 2010) are displayed in the first column followed in the second column by allele assignments proposed by the present comparative analysis.

Glu-B3 allele (Liu et al., 2010)	*Proposed Glu-B3* allele	Cultivar	Genotype										
a	–	'Chinese Spring'	0	636.5	0	0	0	698.6	0	0	729.0	0	0
b	–	'Gabo'	0	636.5	0	0	0	0	0	712.3	0	0	0
c	–	'Halberd'	0	636.5	0	0	696.7	0	0	0	0	791.9	0
d	–	'Pepital'	0	636.5	0	0	0	0	0	0	0	791.9	0
f	–	'Fengmai 27'	0	636.5	0	0	0	0	0	0	0	0	0
g	–	'Capelle-Desprez'	0	636.5	649.8	0	0	0	0	0	0	0	0
h	–	'Neepawa'	0	636.5	0	0	0	0	0	0	0	791.9	0
i	–	'Norin 61'	0	636.5	0	675.5	0	0	0	0	0	791.9	794.8
j	–	'Apollo'	0	636.5	649.8	0	696.7	0	0	0	0	0	0
ab	–	'Nanbu-komugi'	0	636.5	0	675.5	0	0	706.8	0	0	0	0
ac	-	'Thesee'	616.3	0	649.8	675.5	0	0	0	0	0	0	0
ad	-	'Heilo'	616.3	0	0	675.5	0	0	0	0	0	791.9	794.8
b	*b*	'ACA601'	0	636.5	0	0	0	0	0	712.3	0	0	0
b	*b*	'Marquis'	0	636.5	0	0	0	0	0	712.3	0	0	0
c	*c*	'Insignia'	0	636.5	0	0	696.7	0	0	0	0	791.9	0
d/i	*d*	'Orca'	0	636.5	0	0	0	0	0	0	0	791.9	0
h	*h*	'Katepwa'	0	636.5	0	0	0	0	0	0	0	791.9	0
h	*h*	'Klein Flecha'	0	636.5	0	0	0	0	0	0	0	791.9	0
h	*h*	'ProINTA Redomon'	0	636.5	0	0	0	0	0	0	0	791.9	0
h	*h*	'Petrel'	0	636.5	0	0	0	0	0	0	0	791.9	0
h	*c*	'Chopin'	0	636.5	0	0	696.7	0	0	0	0	791.9	0
g	*g*	'Glenlea'	0	636.5	649.8	0	0	0	0	0	0	0	0
f/g/ac	*g*	'Tigre'	0	636.5	649.8	0	0	0	0	0	0	0	0
g/ac	*g*	'Buck Brasil'	0	636.5	649.8	675.5	0	0	0	0	0	0	0
g/ac	*g*	'Klein Proteo'	0	636.5	649.8	675.5	0	0	0	0	0	0	0
g/ac	*ac*	'ACA801'	616.3	0	649.8	675.5	0	0	0	0	0	0	0
j	-	'Klein Martillo'	616.3	0	649.8	0	0	0	0	0	0	0	0
i/ad/d	*ad*	'Opata'	616.3	0	0	675.5	0	0	0	0	0	791.9	794.8
i/d	*ad*	'Etoile de Choisy'	616.3	0	0	675.5	0	0	0	0	0	791.9	0
d/i	*ad*	'Gazul'	616.3	0	0	675.5	696.7	0	0	0	0	791.9	0

described as very frequent in the before cited study, having been observed in 16% of the samples analyzed. In the present work, according to fragment sizes, haplotype *B3-510/570* corresponded to fragments 1-510.0/1-570.1 and 2-616.3/2-675.5 while haplotype *B3-688/691* corresponded to fragments 1-688.0/1-690.5 and 2-791.9. 'Tigre', haplotype *B3-530* of Zhang et al. (2013) corresponded to fragments 1-530.5 and 2-636.5 along with B3-544 to 1-544.0. No correspondence has been established with the Low4 fragments of other studies. Nonetheless, according to segregation and size we can suggest that fragment 4-499.6 corresponded to the *B3-510/570* haplotype, fragments 4-677.8/4-680.6 to the *B3-688/691* haplotype, 4-519.9 to *B3-530* and 4-533.6 to *B3.544*. The fragment 4-610.2 of 'Tigre' could not be assigned to any haplotype previously described.

Some unexpected genotypes were found in the segregations of both generations and could correspond to intra-locus recombination events whose existence, although not very probable, is possible due to the complex organization of the LMW-GS loci formed by several very closely linked genes (Dong et al., 2010). The four individuals of the $F_{4:6}$ generation that presented a distinctive genotype, could have represented intra-locus recombinants which had incorporated fragment 4-610.2 from 'Tigre' into the 'Gazul' genotype. This type of intra-locus recombinant has not previously been described in the literature, most probably because it can only be detected through the analysis of segregating populations. Moreover, usage of a high resolution DNA based method such as the molecular marker system allows for the identification of these recombination types, more difficult to discern by other methodologies. The other three individuals of both populations that presented a distinctive and unique genotype could not be explained as intra-locus recombinants according to the haplotypes described by Zhang et al. (2013).

LMW-GS allele characterization by 1D-SDS-PAGE is often hard to interpret. In reference to the cultivars used as the parental lines of this study, this method has been conclusive in the case of the *Glu-A3* and *Glu-D3* loci, however was ambiguous in the case of the *Glu-B3* locus (Fig. 1). Subunits representing the genetic locus *Glu-B3* were assigned the *d* or *i* allele in the instance of 'Gazul' and the *f* or *g* allele taking into account 'Tigre'. The *Glu-B3* locus comprises the highest variability and most difficult allele determination, the problematical aspect of distinguishing between the *Glu-B3g* and *Glu-B3f* alleles by 1D-SDS-PAGE had previously been described (Wang et al., 2009; Liu et al., 2010). In this sense, 2DE although much more laborious, could entail a higher resolution than 1D-SDS-PAGE. Moreover, there are some alleles (*ab, ac* and *ad*) that can only be identified by 2DE (Liu et al., 2010).

The correlation between the molecular marker approach and the standard cultivars recommended by Liu et al. (2010) for allele determination had previously not been studied. In the present study, a total of 28 cultivars, including the *Glu-B3* standards whose allele composition had been described by Liu et al. (2010) by means of several methodologies (1D-SDS-PAGE, 2DE, PCR and MALDI-TOF) were studied. With the ultimate aim to simplify the method, considering that most fragments were redundant between amplifications, only one primer combination was used. Low2 was selected for enclosing the most external primers to the LMW-GS sequence (See Fig. S1).

The number of fragments amplified was very different between cultivars. This could have reflected pseudogene number differences and not diversity with regard to active genes, in view of the fact that such big differences have at present not been described at a protein level.

Amplified fragments belonged to any of the three different LMW-GS loci. Correct locus assignments can only be achieved by cloning and sequencing all fragments which encompasses a very cumbersome and expensive procedure. We made putative locus assignments, when possible, by comparing fragment sizes to those obtained in a previous study (Zhang et al., 2013), taking into account that slight fragment size differences can be observed between diverse studies.

In a previous study, the molecular marker method had been used to characterize the variability of a set of Chinese bread wheat cultivars, concomitant to determine to some extent the relationship between haplotypes and allele composition (Zhang et al., 2013). Although intricate, some similarities between both studies can be established, i.e. in the present analysis, fragment 2-791.9 differentiated two groups with presence being associated to the *c, d, h, i* and *ad* alleles while absence to the *ab, ac, f, g* and *j* allelic forms. This fragment could possibly be related to the *B3-688* gene described by Zhang et al. (2013) as related to Aroona-*Glu-B3c*, *B3d*, *B3h*, and *B3i*.

Some alleles are extremely difficult to differentiate; whereas *Glu-B3d* vs. *Glu-B3i* or *Glu-B3f* vs. *Glu-B3g* were perfectly discriminated in this comparative analysis, nevertheless, it was not possible to discriminate between *Glu-B3h* vs. *Glu-B3d*. Altogether the *Glu-B3d* allele is difficult to characterize. For instance, cultivar 'Pepital' which is the standard for the *Glu-B3d* allele has also been described as expressing the *Glu-B3i* allelic form, nonetheless, this study has confirmed the constitution as *Glu-B3d*. Two cultivars did not yield expected results, 'Etoile de Choisy' described as *Glu-B3d* or *Glu-B3i* is in this analysis proposed as *Glu-B3ad,* an allelic form very similar to allele *i*. Genotype comparisons identified the *Glu-B3g* allele in the case of 'Tigre' comprising one of the possibilities given by 1D-SDS-PAGE analysis, while the *Glu-B3ad* allele was detected with regard to 'Gazul'. The standard cultivars for the *ab, ac* and *ad* alleles which are only distinguishable by 2DE showed in the present analysis a different genotype, indicative of the need to use a molecular marker system regarding allele characterization as an alternative to 2DE analysis. The discrepancies found in relation to 'Klein Proteo' or 'Buck Brasil' could reflect the fact that our cultivars do not exactly involve the same genetic constitution as that analyzed in other studies. Cultivars with the same denomination holding different genotypes can be found in different laboratories. In this respect, the molecular marker method could constitute a very useful tool to establish and clarify varietal identification.

The results obtained in this study support the use of a molecular marker system for the assignment of *Glu-B3* alleles, in the same way, a possible application for the classification of the *Glu-A3* and *Glu-D3* loci. Further studies contemplating a higher number of cultivars are necessary, taking into account, that only extremely well characterized cultivars using several techniques must be employed. In addition, it is necessary to enhance our knowledge with regard to the effect of LMW-GS in relation to bread quality for a successful application in wheat breeding programs, requiring a better understanding of the genetic constitution and variation of these complex loci.

Acknowledgments

This research was supported by projects AGL2009-09980 and AGL2012-38345 from the Ministry of Science and Innovation of Spain (MICINN). A. Espí received a pre-doctoral fellowship from the MICINN.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jcs.2014.05.008.

References

Ahmad, M., 2000. Molecular marker-assisted selection of HMW glutenin alleles related to wheat bread quality by PCR-generated DNA markers. Theor. Appl. Genet. 101, 892–896.

Békés, F., Kemeny, S., Morell, M., 2006. An integrated approach to predicting end-product quality of wheat. Eur. J. Agron. 25, 155–162.

Dong, L., Zhang, X., Liu, D., Fan, H., Sun, J., Zhang, Z., Qin, H., Li, B., Hao, S., Li, Z., 2010. New insights into the organization, recombination, expression and functional mechanism of low molecular weight glutenin subunit genes in bread wheat. PLoS One 5, e13548.

D'Ovidio, R., Masci, S., 2004. The low-molecular-weight glutenin subunits of wheat gluten. J. Cereal Sci. 39, 321–339.

Gao, L., Ma, W., Chen, J., Wang, K., Li, J., Wang, S., Bekes, F., Appels, R., Yan, Y., 2010. Characterization and comparative analysis of wheat high molecular weight glutenin subunits by SDS-PAGE, RP-HPLC, HPCE, and MALDI-TOF-MS. J. Agric. Food Chem. 58, 2777–2786.

Gupta, R., MacRitchie, F., 1994. Allelic variation at glutenin subunit and gliadin loci, *Glu-1, Glu-3* and *Gli-1* of common wheats. II. biochemical basis of the allelic effects on dough properties. J. Cereal Sci. 19, 19–29.

Gupta, R., Shepherd, K.W., 1990. Two-step one-dimensional SDS-PAGE analysis of LMW subunits ofglutenin. I. variation and genetic control of the subunits in hexaploid wheats. Theor. Appl. Genet. 80, 65–74.

Huang, X., Cloutier, S., 2008. Molecular characterization and genomic organization of low molecular weight glutenin subunit genes at the *Glu-3* loci in hexaploid wheat (*Triticum aestivum* L.). Theor. Appl. Genet. 116, 953–966.

Igrejas, G., Guedes-Pinto, H., Carnide, V., Branlard, G., 1999. The high and low molecular weight glutenin subunits and ω-gliadin composition of bread and durum wheats commonly grown in Portugal. Plant Breed. 118, 297–302.

Ikeda, T., Araki, E., Fujita, Y., Yano, H., 2006. Characterization of low-molecular-weight glutenin subunit genes and their protein products in common wheats. Theor. Appl. Genet. 112, 327–334.

Jackson, E., Morel, M., Sontag Strohm, T., Branlard, G., Metakovsky, E., Redaelli, R., 1996. Proposal for combining the classification systems of alleles of Gli-1 and Glu-3 loci in bread wheat (*Triticum aestivum* L.). J. Genet. Breed. 50, 321–336.

Lan, Q., Feng, B., Xu, Z., Zhao, G., Wang, T., 2013. Molecular cloning and characterization of five novel low molecular weight glutenin subunit genes from Tibetan wheat landraces (*Triticum aestivum* L.). Genet. Resour. Crop Evol. 60, 799–806.

Liu, L., Ikeda, T.M., Branlard, G., Peña, R.J., Rogers, W.J., Lerner, S.E., Kolman, M.A., Xia, X., Wang, L., Ma, W., 2010. Comparison of low molecular weight glutenin subunits identified by SDS-PAGE, 2-DE, MALDI-TOF-MS and PCR in common wheat. BMC Plant Biol. 10, 124.

Long, H., Wei, Y., Yan, Z., Baum, B., Nevo, E., Zheng, Y., 2005. Classification of wheat low-molecular-weight glutenin subunit genes and its chromosome assignment by developing LMW-GS group-specific primers. Theor. Appl. Genet. 111, 1251–1259.

Ma, W., Zhang, W., Gale, K.R., 2003. Multiplex-PCR typing of high molecular weight glutenin alleles in wheat. Euphytica 134, 51–60.

Ma, W., Appels, R., Bekes, F., Larroque, O., Morell, M., Gale, K., 2005. Genetic characterisation of dough rheological properties in a wheat doubled haploid population: additive genetic effects and epistatic interactions. Theor. Appl. Genet. 111, 410–422.

Oury, F., Chiron, H., Faye, A., Gardet, O., Giraud, A., Heumez, E., Rolland, B., Rousset, M., Trottet, M., Charmet, G., 2010. The prediction of bread wheat quality: joint use of the phenotypic information brought by technological tests and the genetic information brought by HMW and LMW glutenin subunits. Euphytica 171, 87–109.

Payne, P.I., Nightingale, M.A., Krattiger, A.F., Holt, L.M., 1987. The relationship between HMW glutenin subunit composition and the bread-making quality of British grown wheat varieties. J. Sci. Food Agric. 40, 51–65.

Saghai-Maroof, M., Soliman, K.M., Jorgensen, R.A., Allard, R.W., 1984. Ribosomal DNA spacer-length polymorphisms in barley: mendelian inheritance, chromosomal location, and population dynamics. Proc. Natl. Acad. Sci. U. S. A. 81, 8014–8018.

Sharma, S., Ram, S., Gupta, R., Sharma, I., 2013. Development of functional marker for distinguishing *Glu-B3b* allele of LMW-GS found in Indian common wheat cultivars. J. Cereal Sci. 57, 245–248.

Shewry, P.R., Halford, N.G., Lafiandra, D., 2003. Genetics of wheat gluten proteins. Adv. Genet. 49, 111–184.

Singh, N., Shepherd, K., Cornish, G., 1991. A simplified SDS-PAGE procedure for separating LMW subunits of glutenin. J. Cereal Sci. 14, 203–208.

Wang, L., Zhao, X., He, Z., Ma, W., Appels, R., Peña, R., Xia, X., 2009. Characterization of low-molecular-weight glutenin subunit *Glu-B3* genes and development of STS markers in common wheat (*Triticum aestivum* L.). Theor. Appl. Genet. 118, 525–539.

Wang, L., Li, G., Peña, R.J., Xia, X., He, Z., 2010. Development of STS markers and establishment of multiplex PCR for *Glu-A3* alleles in common wheat (*Triticum aestivum* L.). J. Cereal Sci. 51, 305–312.

Wrigley, C., 1996. Giant proteins with flour power. Nature 381, 738–739.

Zhang, X., Liu, D., Yang, W., Liu, K., Sun, J., Guo, X., Li, Y., Wang, D., Ling, H., Zhang, A., 2011a. Development of a new marker system for identifying the complex members of the low-molecular-weight glutenin subunit gene family in bread wheat (*Triticum aestivum* L.). Theor. Appl. Genet. 122, 1503–1516.

Zhang, X., Liu, D., Jiang, W., Guo, X., Yang, W., Sun, J., Ling, H., Zhang, A., 2011b. PCR-based isolation and identification of full-length low-molecular-weight glutenin subunit genes in bread wheat (*Triticum aestivum* L.). Theor. Appl. Genet. 123, 1293–1305.

Zhang, X., Jin, H., Zhang, Y., Liu, D., Li, G., Xia, X., He, Z., Zhang, A., 2012. Composition and functional analysis of low-molecular-weight glutenin alleles with Aroona near-isogenic lines of bread wheat. BMC Plant Biol. 12, 243.

Zhang, X., Liu, D., Zhang, J., Jiang, W., Luo, G., Yang, W., Sun, J., Tong, Y., Cui, D., Zhang, A., 2013. Novel insights into the composition, variation, organization, and expression of the low-molecular-weight glutenin subunit gene family in common wheat. J. Exp. Bot. 64, 2027–2040.

Zhao, X., Ma, W., Gale, K., Lei, Z., He, Z., Sun, Q., Xia, X., 2007a. Identification of SNPs and development of functional markers for LMW-GS genes at *Glu-D3* and *Glu-B3* loci in bread wheat (*Triticum aestivum* L.). Mol. Breed. 20, 223–231.

Zhao, X., Xia, X., He, Z., Lei, Z., Appels, R., Yang, Y., Sun, Q., Ma, W., 2007b. Novel DNA variations to characterize low molecular weight glutenin *Glu-D3* genes and develop STS markers in common wheat. Theor. Appl. Genet. 114, 451–460.